

# Every 4-regular 4-uniform hypergraph has a 2-coloring with a free vertex

Michael A. Henning<sup>1,\*</sup> and Anders Yeo<sup>1,2</sup>

<sup>1</sup>Department of Pure and Applied Mathematics  
University of Johannesburg  
Auckland Park, 2006 South Africa  
Email: mahenning@uj.ac.za

<sup>2</sup>Department of Mathematics and Computer Science  
University of Southern Denmark  
Campusvej 55, 5230 Odense M, Denmark  
Email: andersyeo@gmail.com

## Abstract

In this paper, we continue the study of 2-colorings in hypergraphs. A hypergraph is 2-colorable if there is a 2-coloring of the vertices with no monochromatic hyperedge. It is known (see Thomassen [J. Amer. Math. Soc. 5 (1992), 217–229]) that every 4-uniform 4-regular hypergraph is 2-colorable. Our main result in this paper is a strengthening of this result. For this purpose, we define a vertex in a hypergraph  $H$  to be a free vertex in  $H$  if we can 2-color  $V(H) \setminus \{v\}$  such that every hyperedge in  $H$  contains vertices of both colors (where  $v$  has no color). We prove that every 4-uniform 4-regular hypergraph has a free vertex. This proves a known conjecture. Our proofs use a new result on not-all-equal 3-SAT which is also proved in this paper and is of interest in its own right.

**Keywords:** Hypergraphs; Bipartite; 2-Colorable; Transversal; Free vertex; NAE-3-SAT.

**AMS subject classification:** 05C69

---

\*Research supported in part by the South African National Research Foundation and the University of Johannesburg

# 1 Introduction

In this paper, we continue the study of 2-colorings in hypergraphs. We adopt the notation and terminology from [3, 4]. A *hypergraph*  $H = (V, E)$  is a finite set  $V = V(H)$  of elements, called *vertices*, together with a finite multiset  $E = E(H)$  of arbitrary subsets of  $V$ , called *hyperedges* or simply *edges*. A  $k$ -edge in  $H$  is an edge of size  $k$  in  $H$ . The hypergraph  $H$  is  $k$ -uniform if every edge of  $H$  is a  $k$ -edge. The *degree* of a vertex  $v$  in  $H$ , denoted  $d_H(v)$  or simply by  $d(v)$  if  $H$  is clear from the context, is the number of edges of  $H$  which contain  $v$ . The hypergraph  $H$  is  $k$ -regular if every vertex has degree  $k$  in  $H$ . For  $k \geq 2$ , let  $\mathcal{H}_k$  denote the class of all  $k$ -uniform  $k$ -regular hypergraphs. The class  $\mathcal{H}_k$  has been widely studied, both in the context of solving problems on total domination as well as in its own right, see for example [1, 3, 4, 5, 10].

A hypergraph  $H$  is *2-colorable* if there is a 2-coloring of the vertices with no monochromatic hyperedge. Equivalently,  $H$  is 2-colorable if it is *bipartite*; that is, its vertex set can be partitioned into two sets such that every hyperedge intersects both partite sets. Alon and Bregman [1] established the following result.

**Theorem 1** (Alon, Bregman [1]) *Every hypergraph in  $\mathcal{H}_k$  is 2-colorable, provided  $k \geq 8$ .*

Thomassen [12] showed that the Alon-Bregman result in Theorem 1 holds for all  $k \geq 4$ .

**Theorem 2** (Thomassen [12]) *Every hypergraph in  $\mathcal{H}_k$  is 2-colorable, provided  $k \geq 4$ .*

As remarked by Alon and Bregman [1] the result is not true when  $k = 3$ , as may be seen by considering the Fano plane. Sufficient conditions for the existence of a 2-coloring in  $k$ -uniform hypergraphs are given, for example, by Radhakrishnan and Srinivasan [8] and Vishwanathan [13]. For related results, see the papers by Alon and Tarsi [2], Seymour [9] and Thomassen [11].

A set  $X$  of vertices in a hypergraph  $H$  is a *free set* in  $H$  if we can 2-color  $V(H) \setminus X$  such that every edge in  $H$  contains vertices of both colors (where the vertices in  $X$  are not colored). A vertex is a *free vertex* in  $H$  if we can 2-color  $V(H) \setminus \{v\}$  such that every hyperedge in  $H$  contains vertices of both colors (where  $v$  has no color). In [4] it is conjectured that every hypergraph  $H \in \mathcal{H}_k$ , with  $k \geq 4$ , has a free set of size  $k - 3$ . Further, if the conjecture is true, then the bound  $k - 3$  cannot be improved for any  $k \geq 4$ , due to the complete  $k$ -uniform hypergraph of order  $k + 1$ , as such a hypergraph needs two vertices of each color to ensure every edge has vertices of both colors. The conjecture is proved to hold for  $k \in \{5, 6, 7, 8\}$ . The case when  $k = 4$  turned out to be more difficult than the cases when  $k \in \{5, 6, 7, 8\}$  and was conjectured separately in [4].

**Conjecture 1** ([4]) *Every 4-regular 4-uniform hypergraph contains a free vertex.*

## 2 Main Result

Our immediate aim is to prove Conjecture 1. That is, we prove the following result, which is a strengthening of the result of Theorem 2 in the case when  $k = 4$ .

**Theorem 3** *Every 4-regular 4-uniform hypergraph contains a free vertex.*

As remarked earlier, the complete 4-regular 4-uniform hypergraph on five vertices has only one free vertex, and so the result of Theorem 3 cannot be improved in the sense that there exist 4-regular 4-uniform hypergraphs with no free set of size 2. Theorem 3 is also best possible by considering the complement,  $\overline{F_7}$ , of the Fano plane  $F_7$ , where the Fano plane is shown in Figure 1 and where its complement  $\overline{F_7}$  is the hypergraph on the same vertex set  $V(F_7)$  and where  $e$  is a hyperedge in the complement if and only if  $V(F_7) \setminus e$  is a hyperedge in  $F_7$ .

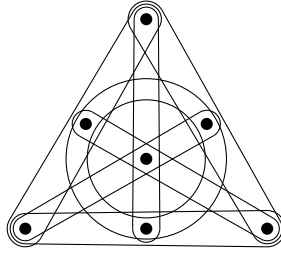


Figure 1: The Fano plane  $F_7$

Our proof of Theorem 3 presented in Section 5 uses a surprising connection with not-all-equal 3-SAT (NAE-3-SAT). We will later prove a result on when NAE-3-SAT is not only satisfiable, but is satisfiable without assigning all variables truth values. This result is of interest in its own right, but requires some further terminology (see Section 3) before describing it in detail. We remark that our resulting NAE-3-SAT result, given by Theorem 4, has also been used by the authors in [6] to solve a conjecture on the so-called fractional disjoint transversal number (which we do not define here). This serves as added motivation of the importance of the NAE-3-SAT result which can be used to solve several seemingly unrelated hypergraph problems that seem difficult to solve using a purely hypergraph approach.

## 3 Terminology and Definitions

For an edge  $e$  in a hypergraph  $H$ , we denote by  $H - e$  the hypergraph obtained from  $H$  by deleting the edge  $e$ . Two vertices  $x$  and  $y$  of  $H$  are *adjacent* if there is an edge  $e$  of  $H$  such that  $\{x, y\} \subseteq e$ . Further,  $x$  and  $y$  are *connected* if there is a sequence  $x = v_0, v_1, v_2, \dots, v_k = y$  of vertices of  $H$  in which  $v_{i-1}$  is adjacent to  $v_i$  for  $i = 1, 2, \dots, k$ . A *connected hypergraph* is a hypergraph in which every pair of vertices are connected. A *component* of a hypergraph

$H$  is a maximal connected subhypergraph of  $H$ . In particular, we note that a component of  $H$  is by definition connected.

A subset  $T$  of vertices in a hypergraph  $H$  is a *transversal* in  $H$  if  $T$  has a nonempty intersection with every edge of  $H$ . In the language of transversals, a vertex  $v$  is a free vertex in a hypergraph  $H$  if  $H$  contains two vertex disjoint transversals, neither of which contain the vertex  $v$ . Transversals in 4-uniform hypergraphs are well studied (see, for example, [4, 7, 10]).

In order to prove Conjecture 1, we use a surprising connection between an instance of not-all-equal 3-SAT (NAE-3-SAT) and a 3-uniform hypergraph. In order to state this connection we require some further terminology.

**Definition 1** *An instance,  $I$ , of 3-SAT contains a set of variables,  $V(I)$ , and a set of clauses,  $C(I)$ . Each clause contains exactly three literals, which are either a variable,  $v \in V(I)$ , or the negation of a variable,  $\bar{v}$ , where  $v \in V(I)$ . A clause,  $c \in C(I)$ , is satisfied if one of the literals in it is true. That is, the clause  $c$  is satisfied if  $v \in V(I)$  belongs to  $c$  and  $v = \text{True}$  or  $\bar{v}$  belongs to  $c$  and  $v = \text{False}$ . The instance  $I$  is satisfied if there is a truth assignment to the variables such that all clauses are satisfied.*

**Definition 2** *An instance of NAE-3-SAT is equivalent to 3-SAT, except that we require all clauses to contain a false literal as well as a true one. A clause that contains both a true and false literal we call nae-satisfiable. If every clause in the instance  $I$  is nae-satisfiable, we say that  $I$  is nae-satisfiable.*

We furthermore need the following definitions.

**Definition 3** *Given an instance  $I$  of NAE-3-SAT, we define the **associated graph**  $G_I$  to be the graph with vertex set  $V(I)$  and where an edge joins two variables in  $G_I$  if they (either in negated or unnegated form) appear in the same clause in  $I$ .*

*Let  $I$  be an instance of NAE-3-SAT. We call the instance  $I$  **connected** if one cannot partition the variables  $V(I)$  into non-empty sets  $V_1$  and  $V_2$  such that no clause contains variables from  $V_1$  and  $V_2$ . In other words, the graph  $G_I$  associated with  $I$  is connected.*

*A **component** of a NAE-3-SAT instance  $I$  is a maximal connected sub-instance of  $I$ . That is, the components of  $I$  correspond precisely to the components of the graph  $G_I$  associated with  $I$ .*

*A variable,  $v \in V(I)$ , is **free** if  $I$  is nae-satisfiable even if we do not assign any truth value to  $v$ . That is, every clause in  $I$  contains a true and a false literal, even without considering literals involving  $v$ .*

*The **degree** of a variable  $v \in V(I)$ , is the number of clauses containing  $v$  or  $\bar{v}$ , and is denoted by  $\deg_I(v)$ . If the instance  $I$  is clear from the context, we simply write  $\deg(v)$  rather than  $\deg_I(v)$ .*

We are now in a position to define a connection between an instance of NAE-3-SAT and a 3-uniform hypergraph as follows.

**Definition 4** *If  $H$  is a 3-uniform hypergraph, we create a NAE-3-SAT instance  $I_H$  as follows. Let  $V(I_H) = V(H)$  and for each edge  $e \in H$  add a clause to  $I_H$  with the same vertices/variables in non-negated form. We call  $I_H$  the NAE-3-SAT instance corresponding to  $H$ . Note that the instance  $I_H$  is nae-satisfiable if and only if  $H$  is bipartite. In fact the partite sets in the bipartition correspond to the truth values true and false.*

Throughout this paper, we use the standard notation  $[k] = \{1, 2, \dots, k\}$ .

## 4 NAE-3-SAT

In this section, we present a key result that we need in order to prove Conjecture 1, namely the following theorem that establishes a fundamental property of NAE-3-SAT in the case when the number of clauses is less than the number of variables. An instance of NAE-3-SAT is non-trivial if it contains at least one variable.

**Theorem 4** *Let  $I$  be a connected non-trivial instance of NAE-3-SAT. If  $|C(I)| < |V(I)|$  and  $\deg_I(v) \leq 3$  for all  $v \in V(I)$ , then  $I$  is nae-satisfiable and contains a free variable.*

**Proof.** Suppose, to the contrary, that the theorem is false and let  $I$  be a counterexample of the theorem with minimum possible  $|C(I)|$ . Let  $C = C(I)$  and  $V = V(I)$ . If  $|C| = 0$ , then  $|V| = 1$  and the theorem holds, and so  $|C| \geq 1$ . We will now show a number of claims which we will use to obtain a contradiction to  $I$  being a counterexample.

**Claim A:**  $\deg(v) \geq 2$  for all  $v \in V$ .

**Proof of Claim A.** Let  $v \in V$  be arbitrary. If  $\deg(v) = 0$ , then  $I$  is not connected as  $|C| \geq 1$ , a contradiction.

Suppose that  $\deg(v) = 1$ , and let  $c \in C$  be the clause containing  $v$ . If  $c$  contains three variables of degree 1, then  $|C| = 1$  and  $|V| = 3$  as  $I$  is connected, and  $I$  is clearly nae-satisfiable and all variables in  $V$  are free, contradicting the fact that  $I$  is a counterexample. Hence, the clause  $c$  has at most two vertices of degree 1. Let  $c$  contain the variables  $v$ ,  $x_1$  and  $x_2$  and let  $I'$  be the instance of NAE-3-SAT obtained by deleting  $v$  and the clause  $c$ .

If  $I'$  is connected, then, by the minimality of  $|C|$ , the instance  $I'$  is nae-satisfiable and contains a free variable. Assigning  $v$  a truth value such that the literal containing  $v$  in the clause  $c$  is of opposite value to the literal containing  $x_1$  or  $x_2$ , the instance  $I$  is nae-satisfiable and contains a free variable, namely the same free variable that belongs to the instance  $I'$ . This contradicts the fact that  $I$  is a counterexample. Therefore,  $I'$  is not connected.

Since  $I$  is connected and  $I'$  is not connected, the instance  $I'$  contains two components, one containing the variable  $x_1$  and the other the variable  $x_2$  that belonged to the clause  $c$ .

Both  $x_1$  and  $x_2$  have degree at most 2 in  $I'$ , while the degree of all other variables in  $I'$  is at most 3. Therefore, both components of  $I'$  satisfy the condition of the theorem and, by the minimality of  $|C|$ , are nae-satisfiable. (We remark that there exists a free variable in both components, but in this case we assign every variable a truth value.) If the literals associated with  $x_1$  and  $x_2$  in the clause  $c$  have the same truth values, then we can reverse the truth value of all variables in one of the components. Hence, we may assume that the literals associated with  $x_1$  and  $x_2$  in the clause  $c$  have different truth values, implying that  $I$  is nae-satisfiable and contains  $v$  as a free variable. Once again, we contradict the fact that  $I$  is a counterexample.

As  $v$  was chosen arbitrarily we have proven Claim A. (□)

**Claim B:** There exists a  $v \in V$ , such that  $\deg(v) = 2$ .

**Proof of Claim B.** If the claim was false, then by Claim A we would have  $\deg(v) = 3$  for all  $v \in V$ , which would imply that  $3|C| = \sum_{v \in V} \deg(v) = 3|V|$ , which is a contradiction to  $|C| < |V|$ . (□)

By Claim B, there exists a variable  $v \in V$  such that  $\deg(v) = 2$ . Let  $c_1$  and  $c_2$  be the clauses containing the variable  $v$  and let  $Q$  contain all variables belonging to  $c_1$  or  $c_2$ .

**Claim C:**  $|Q| \leq 4$ .

**Proof of Claim C.** Suppose, to the contrary, that  $|Q| \geq 5$ . Since the clauses  $c_1$  and  $c_2$  both contain three variables, and  $v$  belongs to both clauses, we note that  $|Q| \leq 5$ . Consequently,  $|Q| = 5$ . Let  $I'$  be the NAE-3-SAT obtained from  $I$  by deleting  $c_1$ ,  $c_2$  and  $v$ .

Suppose that  $I'$  contains four distinct components. In this case, each component of  $I'$  contains a variable from  $Q \setminus \{v\}$ . Possibly, a component of  $I'$  may contain only one variable and no clause. By the minimality of  $|C|$ , the instance  $I'$  is nae-satisfiable. (We remark that there exists a free variable in each of the four components, but in this case we assign every variable a truth value.) We can set the variables in  $Q \setminus \{v\}$  such that both  $c_1$  and  $c_2$  are nae-satisfiable by reversing all truth values in any of the components of  $I'$ , if required, implying that  $I$  is nae-satisfiable and contains  $v$  as a free variable. This contradicts the fact that  $I$  is a counterexample. Therefore,  $I'$  contains at most three distinct components.

Let  $Q \setminus \{v\} = \{q_1, q_2, q_3, q_4\}$ . Renaming variables, if necessary, we may assume that  $q_4$  and one of  $q_1$ ,  $q_2$  or  $q_3$  belong to the same component in  $I'$ . Renaming  $q_1$ ,  $q_2$  and  $q_3$ , if necessary, we may assume that  $q_1$  and  $q_2$  are variables in  $c_1$  and  $q_3$  is a variable in  $c_2$ . If  $v$  is negated in  $c_i$ , then negate all literals in  $c_i$ , for  $i \in [2]$ . This does not change the problem and implies that we may assume, without loss of generality, that  $v$  is not negated in both  $c_1$  and  $c_2$ .

Let  $\ell_i$  be the literal containing  $q_i$  in  $c_1$  for  $i \in [2]$ , and so  $\ell_i \in \{q_i, \bar{q}_i\}$ . Further, let  $\ell_i$  be the literal containing  $q_i$  in  $c_2$  for  $i \in \{3, 4\}$ , and so  $\ell_i \in \{q_i, \bar{q}_i\}$ . Let  $c'$  be a new clause  $\{\ell_1, \ell_2, \bar{\ell}_3\}$  and let  $I'' = I' \cup c'$ . We note that  $I''$  is connected and the degree of all vertices in  $I''$  is at most 3. Further,  $|C(I'')| = |C(I')| + 1 = |C| - 1 < |V| - 1 = |V(I'')|$ . In particular,  $|C(I'')| < |C|$ . By the minimality of  $|C|$ , the instance  $I''$  is nae-satisfiable and contains a

free variable,  $f$ .

If  $f \notin \{q_1, q_2, q_3\}$ , then we can always assign to the variable  $v$  a truth value, such that  $c_1$  and  $c_2$  are nae-satisfiable, as indicated in Table 1, where  $T$  denotes *True* and  $F$  denotes *False* and the cases when  $(\ell_1, \ell_2, \ell_3) \in \{(T, T, F), (F, F, T)\}$  are impossible due to the clause  $c'$ . Further, the free variable,  $f$ , in  $I''$  is also a free variable in  $I$ . Therefore, the instance  $I$  is nae-satisfiable and contains a free variable. This contradicts the fact that  $I$  is a counterexample. Therefore,  $f \in \{q_1, q_2, q_3\}$ .

$(\ell_1, \ell_2, \ell_3)$	$v$	$(\ell_1, \ell_2, \ell_3)$	$v$	$(\ell_1, \ell_2, \ell_3)$	$v$	$(\ell_1, \ell_2, \ell_3)$	$v$
$(T, T, T)$	$F$	$(T, F, T)$	$F$	$(F, T, T)$	$F$	$(F, F, T)$	N/A
$(T, T, F)$	N/A	$(T, F, F)$	$T$	$(F, T, F)$	$T$	$(F, F, F)$	$T$

**Table 1.** Possible assignments of truth values.

Suppose that  $f \in \{q_1, q_2\}$ . Renaming  $q_1$  and  $q_2$ , if necessary, we may assume that  $f = q_1$ . As  $c'$  is nae-satisfiable in  $I''$ , the literals  $\ell_2$  and  $\ell_3$  must have the same truth value. We can therefore assign  $v$  the opposite truth value to these two literals in order to get a nae-satisfiable assignment of  $I$  where the variable  $q_1$  is free. This contradicts the fact that  $I$  is a counterexample. Therefore,  $f = q_3$ .

As  $c'$  is nae-satisfiable in  $I''$  and  $f$  is free in  $I''$ , the literals  $\ell_2$  and  $\ell_3$  must have opposite truth values. We can therefore assign  $v$  the opposite truth value to the literal  $\ell_4$  in order to get a nae-satisfiable assignment of  $I$  where the variable  $q_3$  is free. Once again, this contradicts the fact that  $I$  is a counterexample. This completes all cases and therefore completes the proof of Claim C.  $\square$

**Claim D:**  $|Q| \leq 3$ .

**Proof of Claim D.** Suppose, to the contrary, that  $|Q| \geq 4$ , which by Claim C implies that  $|Q| = 4$ . Therefore, there must exist variables  $q_1, q_2$  and  $q_3$  such that  $c_1$  contains the variables  $v, q_1$  and  $q_2$  and the clause  $c_2$  contains  $v, q_1$  and  $q_3$ . As in the proof of Claim C.3, we may assume, without loss of generality, that  $v$  is not negated in both  $c_1$  and  $c_2$ . Let  $I'$  be the NAE-3-SAT obtained from  $I$  by deleting the two clauses  $c_1$  and  $c_2$ , and deleting the variable  $v$ .

Let  $\ell_i$  be the literal containing  $q_i$  in  $c_1$  for  $i \in [2]$ , and so  $\ell_i \in \{q_i, \bar{q}_i\}$ . Further, let  $\ell_3$  be the literal containing  $q_3$  in  $c_2$ , and so  $\ell_3 \in \{q_3, \bar{q}_3\}$ . Let  $c'$  be a new clause  $\{\ell_1, \ell_2, \bar{\ell}_3\}$  and let  $I'' = I' \cup c'$ . We note that  $I''$  is connected and the degree of all vertices in  $I''$  is at most 3. Further,  $|C(I'')| = |C(I')| + 1 = |C| - 1 < |V| - 1 = |V(I'')|$ . In particular,  $|C(I'')| < |C|$ . By the minimality of  $|C|$ , the instance  $I''$  is nae-satisfiable and contains a free variable,  $f$ .

If  $f \notin \{q_1, q_2, q_3\}$ , then proceeding exactly as in the proof of Claim C, we show that the instance  $I$  is nae-satisfiable and contains a free variable, a contradiction. Therefore,  $f \in \{q_1, q_2, q_3\}$ .

If  $f = q_1$ , then as  $c'$  is nae-satisfiable in  $I''$ , the literals  $\ell_2$  and  $\ell_3$  must have the same truth value. We can therefore assign  $v$  the opposite truth value to these two literals in order

to get a nae-satisfiable assignment of  $I$  where the variable  $q_1$  is free.

If  $f = q_2$ , then as  $c'$  is nae-satisfiable in  $I''$ , the literals  $\ell_1$  and  $\ell_3$  must have the same truth value. We can therefore assign  $v$  the opposite truth value to these two literals in order to get a nae-satisfiable assignment of  $I$  where the variable  $q_2$  is free.

If  $f = q_3$ , then as  $c'$  is nae-satisfiable in  $I''$ , the literals  $\ell_1$  and  $\ell_2$  must have opposite truth values. We can therefore assign  $v$  the opposite truth value to the literal corresponding to  $q_1$  in  $c_2$  in order to get a nae-satisfiable assignment of  $I$  where the variable  $q_3$  is free.

In all the above three cases, the instance  $I$  is nae-satisfiable and contains a free variable, a contradiction. This completes all cases and therefore also the proof of Claim D.  $\square$

By Claim D,  $|Q| \leq 3$ . As every clause contains three variables, this implies that  $|Q| = 3$ . Let  $Q = \{v, q_1, q_2\}$ . Let  $I^*$  be the an instance of NAE-3-SAT with  $V(I^*) = \{v, q_1, q_2\}$  and  $C(I^*) = \{c_1, c_2\}$ .

**Claim E:** The instance  $I^*$  is nae-satisfiable and has a free variable.

**Proof of Claim E.** If at most one literal in  $c_1$  is identical to those in  $c_2$ , then we simply reverse all literals in  $c_1$ . This does not change the problem and now there are at least two literals in  $c_1$  that are identical with those in  $c_2$ . Renaming the variables, if necessary, we may assume that the literal containing  $q_i$  in  $c_1$  and  $c_2$  for  $i \in [2]$  is identical. The variable  $v$  is therefore a free variable as may be seen by assigning opposite truth value to the literal containing  $q_1$  and  $q_2$  in  $c_1$  (and therefore also in  $c_2$ ). Thus,  $I^*$  is nae-satisfiable and has a free variable.  $\square$

By Claim E, the instance  $I^*$  is nae-satisfiable and has a free variable. Therefore,  $I \neq I^*$ , implying that at least one of  $q_1$  and  $q_2$  has degree 3 in  $I$ . There is therefore a clause  $c_3$ , different from  $c_1$  and  $c_2$ , containing  $q_1$  or  $q_2$ . Renaming  $q_1$  and  $q_2$ , if necessary, we may assume that  $c_3$  contains  $q_2$ .

**Claim F:** The clause  $c_3$  does not contain the variable  $q_1$ .

**Proof of Claim F.** Suppose, to the contrary, that  $c_3$  contains  $q_1$ . Let  $q_3$  be the variable in  $c_3$  which is different from  $q_1$  and  $q_2$ . Let  $I''$  be obtained from  $I$  by deleting the three clauses  $c_1$ ,  $c_2$  and  $c_3$ , and deleting the three variables  $v$ ,  $q_1$  and  $q_2$ . We note that  $I''$  is connected and the degree of all vertices in  $I''$  is at most 3. Further,  $|C(I'')| = |C| - 3 < |V| - 3 = |V(I'')|$ . In particular,  $|C(I'')| < |C|$ . By the minimality of  $|C|$ , the instance  $I''$  is nae-satisfiable and contains a free variable. (We remark, however, that here we do not need the fact that there exists a free variable in this case.) By Claim E, it is possible to assign truth values to two of the variables in  $\{v, q_1, q_2\}$  with the third vertex a free variable such that  $c_1$  and  $c_2$  are both nae-satisfiable. At least one of  $q_1$  or  $q_2$  has been assigned a truth value, say  $q_1$ . Let  $\ell_1$  be the literal containing  $q_1$  in  $c_3$ , and let  $\ell_3$  be the literal containing  $q_3$  in  $c_3$ . By reversing all truth values in  $I''$ , if necessary, we can guarantee that the literals  $\ell_1$  and  $\ell_3$  have opposite truth values. Therefore,  $I$  is nae-satisfiable and one of the vertices in  $\{v, q_1, q_2\}$  is free. This is a contradiction to  $I$  being a counterexample.  $\square$



By Claim F, the clause  $c_3$  does not contain the variable  $q_1$ . Let  $I'$  be the NAE-3-SAT obtained from  $I$  by deleting the two clauses  $c_1$  and  $c_2$ , and the variable  $v$ . As in the proof of Claim E, we may assume that there are at least two literals in  $c_1$  that are identical with those in  $c_2$ .

**Claim G:** The variable  $v$  is not free in  $I^*$ .

**Proof of Claim G.** Suppose, to the contrary, that  $v$  is free in  $I^*$ . If the literal containing  $q_1$  in  $c_1$  and  $c_2$  is not identical, then in order for the variable  $v$  to be free in  $I^*$ , the literal containing  $q_2$  in  $c_1$  and  $c_2$  is not identical. This contradicts our assumption that at least two literals in  $c_1$  are identical with those in  $c_2$ . Hence, the literal containing  $q_i$  in  $c_1$  and  $c_2$  for  $i \in [2]$  is identical.

If in  $c_1$  and  $c_2$  exactly one of  $q_1$  and  $q_2$  is negated, then let  $c'_3$  be a new clause obtained from  $c_3$  by replacing the literal  $q_2$  with  $q_1$  or  $\bar{q}_2$  with  $\bar{q}_1$ . If in  $c_1$  and  $c_2$  either both or none of  $q_1$  and  $q_2$  are negated, then let  $c'_3$  be a new clause obtained from  $c_3$  by replacing the literal  $q_2$  with  $\bar{q}_1$  or  $\bar{q}_2$  with  $q_1$ . Let  $I''$  be the instance obtained from  $I'$  by deleting the clause  $c_3$  and the variable  $q_2$ , and adding the clause  $c'_3$ . We note that  $I''$  is connected and the degree of all vertices in  $I''$  is at most 3. Further,  $|C(I'')| = |C| - 2 < |V| - 2 = |V(I'')|$ . In particular,  $|C(I'')| < |C|$ .

By the minimality of  $|C|$ , the instance  $I''$  is nae-satisfiable and contains a free variable. (We remark, however, that here we do not need the fact that there exists a free variable in  $I''$  in this case.) If in  $c_1$  and  $c_2$  exactly one of  $q_1$  and  $q_2$  is negated, then we let  $q_2$  have the same truth value as  $q_1$ ; otherwise, we let  $q_2$  have the opposite truth value of  $q_1$ . This implies that  $I$  is nae-satisfiable even without assigning  $v$  a value. Therefore, the variable  $v$  is free and  $I$  is not a counterexample to the theorem, a contradiction.  $\square$

By Claim G, the variable  $v$  is not free in  $I^*$ . By our earlier assumptions, there are at least two literals in  $c_1$  that are identical with those in  $c_2$ . If the literal containing  $q_i$  in  $c_1$  and  $c_2$  is identical for  $i \in [2]$ , then  $v$  is free in  $I^*$ , a contradiction. This implies that if the literal containing  $q_i$  in  $c_1$  and  $c_2$  is not identical, then the literal containing  $q_{3-i}$  in  $c_1$  and  $c_2$  is identical for  $i \in [2]$ . Further, the literal containing  $v$  in  $c_1$  and  $c_2$  is identical for  $i \in [2]$ .

Let  $c'_3$  be a new clause obtained from  $c_3$  by replacing the literal  $q_2$  with  $q_1$  or  $\bar{q}_2$  with  $\bar{q}_1$ . Let  $I''$  be the instance obtained from  $I'$  by deleting the clause  $c_3$  and the variable  $q_2$ , and adding the clause  $c'_3$ . We note that  $I''$  is connected and the degree of all vertices in  $I''$  is at most 3. Further,  $|C(I'')| = |C| - 2 < |V| - 2 = |V(I'')|$ . In particular,  $|C(I'')| < |C|$ . By the minimality of  $|C|$ , the instance  $I''$  is nae-satisfiable and contains a free variable.

Suppose that  $q_1$  is free in  $I''$ . In this case, we can assign values to  $v$  and  $q_2$  if  $q_1$  is free in  $I^*$  and to  $v$  and  $q_1$  if  $q_2$  is free in  $I^*$ , such that  $c_1$  and  $c_2$  are nae-satisfiable. Therefore,  $I$  is nae-satisfiable and contains  $q_1$  or  $q_2$  as a free variable. This is a contradiction to  $I$  being a counterexample. Hence,  $q_1$  is not free in  $I''$ .

Since  $I''$  contains a free variable and  $q_1$  is not free in  $I''$ , some variable,  $w$ , different from  $q_1$  is a free variable in  $I''$ . We now assign  $q_2$  the same truth value as  $q_1$  and we assign  $v$  the opposite truth value to the literal corresponding to  $q_1$  in  $c_1$  (or  $c_2$ ). With this truth

assignment, both  $c_1$  and  $c_2$  are nae-satisfiable, noting that the literal containing  $v$  in  $c_1$  and  $c_2$  is identical for  $i \in [2]$ . Therefore,  $I$  is nae-satisfiable and contains  $w$  as a free variable. Once again, this is a contradiction to  $I$  being a counterexample, which completes the proof of Theorem 4.  $\square$

We remark that the result of Theorem 4 is best possible in the following sense.

**Proposition 1** *For any  $s \geq 1$ , there exists a non-trivial connected instance  $I$  of NAE-3-SAT with  $3s$  variables satisfying  $0 \leq |C(I)| < |V(I)|$  and  $\deg_I(v) \leq 3$  for all  $v \in V(I)$  such that  $I$  is nae-satisfiable and contains exactly one free variable.*

**Proof.** Let  $s \geq 1$  and let  $I$  be an instance of NAE-3-SAT with variables  $V(I) = \{v_i^j \mid i \in [s] \text{ and } j \in [3]\}$  and clauses  $C(I) = C_1 \cup C_2$ , where

$$\begin{aligned} C_1 &= \{(v_i^1, v_i^2, v_i^3), (\bar{v}_i^1, v_i^2, v_i^3) \mid i \in [s]\} \\ C_2 &= \{(v_i^1, v_{i+1}^2, \bar{v}_{i+1}^3) \mid i \in [s-1]\}. \end{aligned}$$

By construction,  $I$  is connected and the degree of all vertices in  $I$  is at most 3. Further,  $|C(I)| = 3s - 1 = |V(I)| - 1$ . We will now show that the only free variable in  $I$  is  $v_s^1$ . Due to the clauses in  $C_1$  we note that  $v_i^2$  and  $v_i^3$  must be assigned opposite truth values in any nae-satisfiable truth assignment for all  $i \in [s]$ . For every  $i \in [s-1]$ , we note that  $v_{i+1}^2$  and  $\bar{v}_{i+1}^3$  have the same truth value since  $v_{i+1}^2$  and  $v_{i+1}^3$  have opposite truth values. This implies that  $v_i^1$  must be assigned the opposite truth value to  $v_{i+1}^2$  and  $\bar{v}_{i+1}^3$ . This is true for all  $i \in [s-1]$ , which implies that  $v_s^1$  is the only variable that can be free. It is not difficult to see that  $v_s^1$  is free and this also follows from Theorem 4, noting that none of the other variables are free.  $\square$

## 5 Proof of Theorem 3

Using Theorem 4, we prove Theorem 3. First, we present the following lemma.

**Lemma 5** *Let  $H$  be a connected 3-uniform hypergraph with no isolated vertex. If  $H$  has fewer edges than vertices and has maximum degree at most 3, then  $H$  contains at least two free vertices.*

**Proof.** Let  $H$  be a connected 3-uniform hypergraph with no isolated vertex. Suppose that  $H$  has fewer edges than vertices and has maximum degree at most 3. Let  $I_H$  be the NAE-3-SAT instance corresponding to  $H$ . By Theorem 4, the instance  $I_H$  is nae-satisfiable and has a free vertex, say  $v$ . Assigning color 1 to true variables and color 2 to false variables we obtain a 2-coloring  $\mathcal{C}$  of  $H$  where  $v$  has no color and all hyperedges of  $H$  contain vertices of both colors. Let  $E_v$  be all edges in  $H$  containing the vertex  $v$ . Since  $H$  has no isolated vertex, we note that  $|E_v| = d_H(v) \geq 1$ .

We say that a vertex in  $H$  is  $\mathcal{C}$ -fixed if in some edge in  $E(H) \setminus E_v$  it is the only vertex of its color in the 2-coloring  $\mathcal{C}$ . We note that every edge of  $H$  is a 3-edge, and every edge in  $E(H) \setminus E_v$  contains vertices of both colors in  $\mathcal{C}$ . Thus, in every edge in  $E(H) \setminus E_v$  there is a vertex whose color is unique in that edge. Thus, every edge in  $E(H) \setminus E_v$  gives rise to exactly one vertex that is  $\mathcal{C}$ -fixed. Therefore, there are at most  $|E(H) \setminus E_v|$  vertices in  $H$  that are  $\mathcal{C}$ -fixed.

By supposition,  $|E(H)| < |V(H)|$ . Hence,  $|E(H) \setminus E_v| \leq |E(H)| - |E_v| \leq (|V(H)| - 1) - d_H(v) \leq |V(H)| - 2$ , implying that at least two vertices in  $H$  are not  $\mathcal{C}$ -fixed. Clearly, the vertex  $v$  is not  $\mathcal{C}$ -fixed. Let  $u$  be a vertex different from  $v$  that is not  $\mathcal{C}$ -fixed. Renaming colors if necessary, we may assume that  $u$  has color 1. Thus, every edge in  $E(H) \setminus E_v$  that contains  $u$  contains another vertex of color 1 and one vertex of color 2. Let  $\mathcal{C}'$  be the coloring obtained from  $\mathcal{C}$  by removing the color 1 from  $u$  and assigning color 1 to  $v$ . Since  $\mathcal{C}$  is a 2-coloring of  $H$ , so too is  $\mathcal{C}'$  a 2-coloring of  $H$ . However, in the 2-coloring  $\mathcal{C}'$  the vertex  $u$  is a free vertex. Thus,  $H$  has at least two free vertices, namely  $u$  and  $v$ .  $\square$

We are now in a position to prove Theorem 3. Recall its statement.

**Theorem 3.** *Every 4-regular 4-uniform hypergraph contains a free vertex.*

**Proof of Theorem 3.** We may assume that  $H$  is connected as otherwise we consider each component of  $H$  separately. By Thomassen's Theorem 2, there exists a 2-coloring,  $\mathcal{C}$ , of  $H$  such that no edge of  $H$  is monochromatic. Analogously to the proof of Lemma 5, we call a vertex  $\mathcal{C}$ -fixed if in some edge in  $E(H)$  it is the only vertex of its color in  $\mathcal{C}$ . If some vertex is not  $\mathcal{C}$ -fixed, then it is a free vertex, as we can simply uncolor it. Therefore, we may assume that all vertices in  $H$  are  $\mathcal{C}$ -fixed, for otherwise the desired result follows.

For every edge  $e \in E(H)$ , let  $v^*(e)$  be the vertex of unique color in  $e$ , if such a vertex exists. By assumption, all vertices in  $H$  are  $\mathcal{C}$ -fixed, implying that for every vertex  $u \in V(H)$  we have  $u = v^*(e)$  for some edge  $e$  in  $H$ . Since  $H$  is a 4-regular 4-uniform hypergraph, we note that  $|V(H)| = |E(H)|$ . Thus since every vertex in  $H$  is  $\mathcal{C}$ -fixed, this implies that for every edge  $e \in E(H)$ , the vertex  $v^*(e)$  exists. Further, if  $e$  and  $e'$  are distinct edges, then  $v^*(e) \neq v^*(e')$ . This in turn implies that for every vertex  $u \in V(H)$  there is a unique edge,  $e^*(u)$ , such that  $v^*(e^*(u)) = u$ .

Let  $V_1$  be the set of all vertices of color 1 in  $\mathcal{C}$  and let  $V_2$  be the set of all vertices of color 2 in  $\mathcal{C}$ . For each vertex  $u \in V_1$ , the edge  $e^*(u)$  contains three vertices in  $V_2$ , while for each vertex  $v \in V_2$ , the edge  $e^*(v)$  contains one vertex in  $V_2$ . Thus, the sum of the degrees of the vertices in  $V_2$  is  $3|V_1| + |V_2|$ , implying that the average degree of a vertex in  $V_2$  is  $(3|V_1| + |V_2|)/|V_2|$ . Since  $H$  is 4-regular, this value has to be 4, which implies that  $|V_1| = |V_2|$ .

Let  $H_1^*$  be the hypergraph with vertex set  $V(H_1^*) = V_1$  and with edge set defined as follows: for every vertex  $u \in V_2$  add the edge  $e_u = (e^*(u) \setminus \{u\})$  to  $H_1^*$ . We note that each vertex  $v \in V_1$  belongs to one edge  $e^*(v)$  of  $H$  and to three edges of the type  $e^*(u)$  where  $u \in V_2$ . Thus, by construction,  $H_1^*$  is a 3-regular 3-uniform hypergraph. Analogously, we

define  $H_2^*$  be the hypergraph with vertex set  $V(H_2^*) = V_2$  and with edge set defined as follows: for every  $u \in V_1$  add the edge  $e_u = (e^*(v) \setminus \{v\})$  to  $H_2^*$ . By construction,  $H_2^*$  is a 3-regular 3-uniform hypergraph. Let  $C_1^1, \dots, C_{\ell_1}^1$  be the components of  $H_1^*$  where  $\ell_1 \geq 1$ , and let  $C_1^2, \dots, C_{\ell_2}^2$  be the components of  $H_2^*$  where  $\ell_2 \geq 1$ . Let  $i_1 = 1$ .

Let  $u_1 \in V(C_{i_1}^1)$  and let  $i_2$  be defined such that  $e_{u_1}$  is an edge in  $C_{i_2}^2$ . We note that  $C_{i_2}^2 - e_{u_1}$  contains at most three components. Further, every component of  $C_{i_2}^2 - e_{u_1}$  has fewer edges than vertices as the degrees of its vertices are at most 3 and it contains a vertex of degree at most 2, namely a vertex contained in the deleted edge  $e_{u_1}$ . Therefore applying Lemma 5 to each component of  $C_{i_2}^2 - e_{u_1}$ , we obtain a 2-coloring of the component that contains a free vertex. Combining these 2-colorings in each component, produces a 2-coloring of  $C_{i_2}^2 - e_{u_1}$  that contains at least one free vertex. Let  $u_2$  be a free vertex of  $C_{i_2}^2 - e_{u_1}$ .

Let  $i_3$  be defined such that  $e_{u_2}$  is an edge in  $C_{i_3}^1$ . As before, applying Lemma 5 to each component of  $C_{i_3}^1 - e_{u_2}$ , we obtain a 2-coloring with a free vertex, say  $u_3$ .

Continuing the above process we obtain a sequence  $i_1, i_2, i_3, i_4, \dots$ . As there are finitely many components of  $H_1^*$  and  $H_2^*$ , we note that there must exist integers  $\ell$  and  $k$ , such that  $i_\ell, i_{\ell+2}, i_{\ell+4}, \dots, i_{k-2}$  are all distinct and  $i_{\ell+1}, i_{\ell+3}, i_{\ell+5}, \dots, i_{k-1}$  are all distinct and  $i_\ell = i_k$ .

Renaming components if necessary, we may assume we had started with  $u_\ell$  instead of  $u_1$  and we may assume  $\ell = 1$  and  $i_1 = 1, i_2 = 1, i_3 = 2, i_4 = 2, i_5 = 3, \dots, i_{k-2} = (k-1)/2, i_{k-1} = (k-1)/2$  and  $i_k = 1$ . By Lemma 5, the hypergraph  $C_{i_k}^1 - e_{u_{k-1}} = C_1^1 - e_{u_{k-1}}$  contains at least two free vertices. We may without loss of generality assume that the free vertex  $u_k$  of  $C_1^1 - e_{u_{k-1}}$  was chosen to be different from  $u_1$ . We now define the sets  $V'$  and  $E'$  by  $V' = \{u_2, \dots, u_{k-1}, u_k\}$  and  $E' = \{e_{u_1}, e_{u_2}, \dots, e_{u_{k-1}}\}$ . Further we define

$$V^* = \bigcup_{i=1}^{\frac{k-1}{2}} V(C_i^1) \cup V(C_i^2) \quad \text{and} \quad E^* = \bigcup_{i=1}^{\frac{k-1}{2}} E(C_i^1) \cup E(C_i^2).$$

If  $e_u$  is an edge in  $H_1^*$  or  $H_2^*$ , then let  $(e_u)^H$  be the 4-edge containing the vertices  $V(e_u) \cup \{u\}$ ; that is,  $(e_u)^H$  is the original 4-edge in  $H$  that gave rise to  $e_u$ . We now define

$$E^{**} = \{(e)^H \mid e \in E^*\} \quad \text{and} \quad E'' = \{(e)^H \mid e \in E'\}.$$

We note that  $E' \subset E^*$  and every edge in  $E^*$  is a 3-edge. Further,  $E'' \subset E^{**}$  and every edge in  $E^{**}$  is a 4-edge. Considering the 2-colorings we obtained above we can 2-color the vertices of  $V^* \setminus V'$  such that all edges of  $E^* \setminus E'$  (and therefore also all edges of  $E^{**} \setminus E''$ ) contain vertices of both colors. Interchanging the colors of all vertices in  $C_1^1$  if necessary (by recoloring vertices of color  $i$  with color  $3-i$  for  $i \in [2]$  in the original 2-coloring of  $C_1^1$ ), we may assume that  $(e_{u_1})^H$  also contains vertices of both colors, noting that  $u_1 \neq u_k$ . Coloring the vertex  $u_i$  we can make sure that the edge  $(e_{u_i})^H$  contains vertices of both colors, for each  $i \in [k-1] \setminus \{1\}$ . We have now 2-colored all vertices in  $V^*$  except for the vertex  $u_k$  (which is still uncolored) such that all edges in  $E^{**}$  contain vertices of both colors. Let  $\mathcal{C}^*$  denote the resulting 2-coloring of the vertices of  $V^*$ . If  $V^* = V(H)$ , then  $\mathcal{C}^*$  is a 2-coloring

of  $H$  with a free vertex, and we are therefore done and the desired result follows. Hence, we may assume that  $V^* \neq V(H)$ .

If the  $|V^*|$  edges in  $\{e_v \mid v \in V^*\}$  are exactly the edges in  $E^*$ , then  $H$  would not be connected, a contradiction. Therefore, there must be a vertex  $v \in V^*$  where  $e_v \notin E^*$ . Hence,  $e_v \in C_a^b$  for some  $a > (k-1)/2$  and  $b \in [2]$ . Applying Lemma 5 to each component of  $C_a^b - e_v$ , we obtain analogously as before a 2-coloring of  $C_a^b - e_v$  with a free vertex, say  $x_a^b$ . If the vertex  $v$  was already colored in  $\mathcal{C}^*$ , then interchange all colors in the 2-coloring  $\mathcal{C}^*$ , if necessary, in order to guarantee that the edge  $(e_v)^H$  contains vertices of both colors. If  $v$  was not colored, then color it such that  $(e_v)^H$  contains vertices of both colors. In both cases, we produce a 2-coloring of the vertices of  $V^* \cup V(C_a^b)$  with the vertex  $x_a^b$  uncolored (and possibly also the vertex  $u_k$  uncolored if  $v \neq u_k$ ) such that all edges in  $E^{**} \cup \{(e)^H \mid e \in E(C_a^b)\}$  contain vertices of both colors.

Repeating the above process (with the new  $V^*$  being  $V^* \cup V(C_a^b)$  and the new  $E^{**}$  being  $E^{**} \cup \{(e)^H \mid e \in E(C_a^b)\}$ ), we will eventually have  $V^* = V(H)$  and produce a 2-coloring of  $H$  with a free vertex. This completes the proof of Theorem 3.  $\square$

## 6 Closing Remarks

In this paper, we establish a surprising connection between NAE-3-SAT and 2-coloring of hypergraphs. We prove that every connected non-trivial instance of NAE-3-SAT with maximum degree 3 is nae-satisfiable (and contains a free variable) if the number of clauses is less than the number of variables. Using this property, we strengthen a beautiful result due to Carsten Thomassen [12] that every 4-regular 4-uniform hypergraph is 2-colorable, which itself is a strengthening of a powerful result due to Alon and Bregman [1].

As remarked earlier, our result (see Theorem 3) is best possible in the sense that there exist 4-regular 4-uniform hypergraphs with only free vertex; that is, every free set in such a hypergraph has size 1. We believe that every connected 4-regular 4-uniform hypergraph with sufficiently large order contains a free set of size 2. Due to the complement of the Fano plane the order of such a hypergraph is more than 7. It is possible that every connected 4-regular 4-uniform hypergraph of order at least 8 contains a free set of size 2. We further suspect that every connected 4-regular 4-uniform hypergraph with sufficiently large order  $n$  contains a free set of size  $C \times n$ , where  $C > 0$  is some constant.

As remarked previously, we have subsequently used our NAE-3-SAT property, given by Theorem 4, to solve other seemingly unrelated hypergraph conjectures (such as a conjecture on the fractional disjoint transversal number) that seem difficult to solve using a purely hypergraph approach.

An interesting line of future research would be to determine for larger values of  $k \geq 4$  which connected non-trivial NAE- $k$ -SAT instances are nae-satisfiable given the maximum degree, number of variables and number of clauses, and to apply such results to solve open problems and conjectures related to  $k$ -uniform hypergraphs. We believe this would be a very interesting avenue of research to explore.

## References

- [1] N. Alon and Z. Bregman, Every 8-uniform 8-regular hypergraph is 2-colorable. *Graphs Combin.* **4** (1988), 303–306.
- [2] N. Alon and M. Tarsi, Colorings and orientations of graphs. *Combinatorica* **12** (1992) 124–134.
- [3] M. A. Henning and A. Yeo, 2-Colorings in  $k$ -regular  $k$ -uniform hypergraphs. *European J. Combin.* **34** (2013), 1192–1202.
- [4] M. A. Henning and A. Yeo, On 2-colorings of hypergraphs. *J. Graph Theory* **80**(2) (2015), 112–135.
- [5] M. A. Henning and A. Yeo, Transversals in 4-uniform hypergraph. *Electronic J. Combin.* **23**(3) (2016), #P3.50.
- [6] M. A. Henning and A. Yeo, Fractional disjoint transversals in hypergraphs, manuscript, 2016.
- [7] F. C. Lai and G. J. Chang, An upper bound for the transversal numbers of 4-uniform hypergraphs. *J. Combin. Theory Ser. B* **50** (1990), 129–133.
- [8] J. Radhakrishnan and A. Srinivasan, Improved bounds and algorithms for hypergraph 2-coloring. *Random Struct. Algorithms* **16** (2000), 4–32.
- [9] P. D. Seymour, On the two coloring of hypergraphs. *Quart. J. Math. Oxford Ser.* **25** (1974), 303–312.
- [10] S. Thomassé and A. Yeo, Total domination of graphs and small transversals of hypergraphs. *Combinatorica* **27** (2007), 473–487.
- [11] C. Thomassen, Disjoint cycles in digraphs. *Combinatorica* **3** (1983), 393–396.
- [12] C. Thomassen, The even cycle problem for directed graphs. *J. Amer. Math. Soc.* **5** (1992), 217–229.
- [13] S. Vishwanathan, On 2-coloring certain  $k$ -uniform hypergraphs. *J. Combin. Theory Ser. A* **101** (2003), 168–172.